

AD-A150 589

ON LIMIT OF THE LARGEST EIGENVALUE OF THE LARGE
DIMENSIONAL SAMPLE COVARI. (U) PITTSBURGH UNIV PA
CENTER FOR MULTIVARIATE ANALYSIS Y Q VIN ET AL OCT 84
UNCLASSIFIED TR-84-44 AFOSR-TR-85-0006 F49620-82-K-0001 F/G 12/1

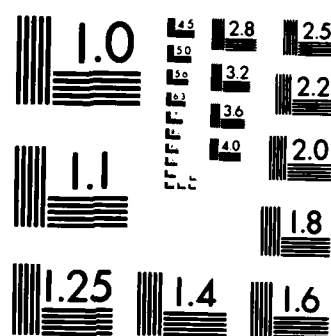
1/1

NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

3

AD-A150 589

ON LIMIT OF THE LARGEST EIGENVALUE
OF THE LARGE DIMENSIONAL SAMPLE
COVARIANCE MATRIX

Y. Q. Yin¹, Z. D. Bai¹,
and
P. R. Krishnaiah²

F49620-82-X-0001

Center for Multivariate Analysis
University of Pittsburgh

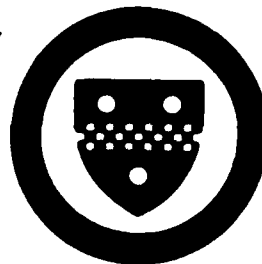
DTIC FILE COPY

DTIC
ELECTE
FEB 22 1985

S

B

D



Approved for public release;
distribution unlimited.

ON LIMIT OF THE LARGEST EIGENVALUE
OF THE LARGE DIMENSIONAL SAMPLE
COVARIANCE MATRIX

Y. Q. Yin¹, Z. D. Bai¹,
and
P. R. Krishnaiah²

F49620-82-X-0001

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
October 1984 NOTICE OF TRANSFER OF RIGHTS
This technical report is the property of the AFSC.
approved for release by AFSC on 10-12-84.
Distribution is unlimited.
MATTHEW J. KEEPER
Chief, Technical Information Division

Technical Report No. 84-44

Center for Multivariate Analysis
515 Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

DTIC
ELECTE
S FEB 22 1985 D
B

¹Y. Q. Yin and Z. D. Bai are on leave of absence from China University of Science and Technology.

²The work of this author is supported by the Air Force Office of Scientific Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

ON LIMIT OF THE LARGEST EIGENVALUE OF THE
LARGE DIMENSIONAL SAMPLE COVARIANCE MATRIX

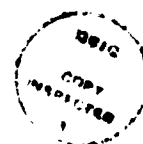
Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah

ABSTRACT

In this paper, the authors showed that the largest eigenvalue of the sample covariance matrix tends to a limit under certain conditions when both the number of variables and the sample size tend to infinity. The above result is proved under the mild restriction that the fourth moment of the elements of the sample sums of squares and cross products (SP) matrix exist.

Keywords: Largest eigenvalue, sample covariance matrix, large dimensional random matrices, limit.

Accession For	
NTIS	<input checked="checked" type="checkbox"/>
DTIC	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Distribution/	
Availability Codes	
Avail and/or	
Special	
Dist	
A-1	



1. INTRODUCTION

The distribution of the largest eigenvalue of the sample covariance matrix is useful in certain problems of inference in the area of multivariate analysis. For example, it is useful in testing the hypothesis that the eigenvalues of the covariance matrix are equal to a specified value. Geman (1980) showed that the largest eigenvalue of $A = WW'/n$ tends to $(1 + \sqrt{y})^2$ almost surely when $\lim (p/n) = y$, $W = (w_{ij})$ and w_{ij} 's are distributed independently with mean zero and variance one. In proving the above result, Geman assumed that $E|w_{11}|^n \leq n^{\alpha n}$ for $n = 1, 2, \dots$ and a positive constant α . Jonsson (1983) announced the above result under the weaker condition that $E(|w_{11}|^7) < \infty$ by using a "truncation" method. Recently, Silverstein (1984) proved the same result under the condition that $E|w_{11}|^{6+\varepsilon} < \infty$ where $\varepsilon > 0$ is arbitrary. In the present paper, we prove the above result under a much weaker condition that $E|w_{11}|^4 < \infty$.

2. PRELIMINARIES

The following results are needed in the sequel:

Lemma 2.1. For any $\delta > 0$, we have

$$\binom{k}{r} \binom{k}{\ell-r} \binom{2k-\ell}{\ell} \delta^{k-\ell} \leq (1 + \sqrt{\delta})^{2k} \binom{k}{\ell} \binom{2\ell}{2r}$$

for all $1 \leq r \leq \ell \leq k$, where $\binom{n}{m} = n!/m!(n-m)!$.

Proof. Let

$$I(r) = \binom{k}{r} \binom{k}{\ell-r} / \binom{2\ell}{2r} \text{ for } r = 0, 1, \dots, \ell.$$

Then, we have for $r = 0, 1, 2, \dots, \ell - 1$,

$$\frac{I(r+1)}{I(r)} = \frac{(2r+1)(k-r)}{[2(\ell-r)-1][k-(\ell-r)+1]}.$$

But

$$\begin{aligned} & (2r+1)(k-r) - [2(\ell-r)-1][k-(\ell-r)+1] \\ &= (2r-\ell+1)(2k-2\ell+1) \geq 0 \text{ iff } 2r \geq \ell-1. \end{aligned}$$

So,

$$\frac{I(r+1)}{I(r)} \geq 1 \text{ iff } 2r \geq \ell-1.$$

Hence $I(r)$ has its maximum at $I(0) = I(\ell)$, and so

$$\begin{aligned} & \left(\binom{k}{r} \binom{k}{\ell-r} \binom{2k-\ell}{\ell} / \binom{k}{\ell} \binom{2\ell}{2r} \right) \delta^{k-\ell} \leq \binom{2k-\ell}{\ell} \delta^{k-\ell} \leq \sum_{\ell=0}^k \binom{2k-\ell}{\ell} \delta^{k-\ell} \\ &= \sum_{\ell=0}^k \binom{k+\ell}{\ell} \delta^{\ell} \leq \sum_{\ell=0}^{2k} \binom{2k}{\ell} \delta^{\ell/2} = (1 + \sqrt{\delta})^{2k}. \end{aligned}$$

In studying strong limit properties of random matrix, the techniques of truncation and centralization play an important role.

Lemma 2.2. (Truncation Lemma). Let r be a number belonging to the interval $[\frac{1}{2}, 2]$ and let $\{w_{ij}, i, j = 1, 2, \dots\}$ be a collection of iid random variables with $E w_{11} = 0$ and $E |w_{11}|^{2/r} < \infty$. For each n , define

W_n to be $p \times n$ matrix whose (i,j) th entry is w_{ij} , where $p = p(n)$ is such that $p/n \rightarrow y \in (0, \infty)$ as $n \rightarrow \infty$. Then there exists a sequence of positive numbers $\delta = \delta_n$ such that

- (a) $\delta \rightarrow 0$ as $n \rightarrow \infty$,
- (b) The convergence rate of δ can be slower than any preassigned rate,
- (c) $P(W_n \neq \hat{W}_n, \text{i.o.}) = 0$.

where \hat{W} is the $p \times n$ matrix whose (i,j) -th entry is $\hat{w}_{ij} = w_{ij} I[|w_{ij}| < \delta n^r]$ and I_A denotes the indicator of the set A .

Proof. Since $E|w_{11}|^{2/r} < \infty$, we have for any $\epsilon > 0$

$$\sum_{m=1}^{\infty} 2^{2m} P(|w_{11}| \geq \epsilon 2^{mr}) < \infty.$$

Because of the arbitrariness of ϵ in the above inequality, there exists a sequence of positive numbers $\epsilon = \epsilon_m$ such that

- (1) $\epsilon_m \rightarrow 0$, when $m \rightarrow \infty$,
- (2) The rate of convergence is slower than any preassigned rate,
- (3) $\sum_{m=1}^{\infty} 2^{2m} P(|w_{11}| \geq \epsilon_m 2^{mr}) < \infty$.

Define $\delta = \delta_n = 2\epsilon_m$ for $2^{m-1} \leq n < 2^m$. It is obvious that such a sequence of δ satisfies the requirements (a) and (b). Define \hat{W}_n with this δ . Because $\frac{p}{n} \rightarrow y$, $0 < y < \infty$, we have $p \leq 2yn$ when n is large enough. Thus

$$\begin{aligned} & P(W_n \neq \hat{W}_n, \text{i.o.}) \\ & \leq \lim_{k \rightarrow \infty} \sum_{m=k}^{\infty} P\left(\bigcup_{2^{m-1} \leq n < 2^m} \bigcup_{i=1}^p \bigcup_{j=1}^n (|w_{ij}| \geq \delta n^r)\right) \\ & \leq \lim_{k \rightarrow \infty} \sum_{m=k}^{\infty} P\left(\bigcup_{2^{m-1} \leq n < 2^m} \bigcup_{i=1}^{2y2^m} \bigcup_{j=1}^{2^m} (|w_{ij}| \geq \epsilon_m 2^{mr})\right) \\ & = \lim_{k \rightarrow \infty} \sum_{m=k}^{\infty} P\left(\bigcup_{i=1}^{2y2^m} \bigcup_{j=1}^{2^m} (|w_{ij}| \geq \epsilon_m 2^{mr})\right) \\ & \leq \lim_{k \rightarrow \infty} 2y \sum_{m=k}^{\infty} 2^{2m} P(|w_{11}| \geq \epsilon_m 2^{mr}) = 0 \end{aligned}$$

which completes the proof.

Lemma 2.3. (Centralization Lemma) Under the assumptions of Lemma 2.2, we have

$$|\tilde{\lambda}_{\max}(n) - \hat{\lambda}_{\max}(n)| \rightarrow 0 \text{ a.s.} \quad (2.1)$$

where $\tilde{\lambda}_{\max}(n)$ and $\hat{\lambda}_{\max}(n)$ are the largest eigenvalues of $\frac{1}{n^{2r}} \tilde{W}_n \tilde{W}_n'$ and $\frac{1}{n^{2r}} \hat{W}_n \hat{W}_n'$, respectively, and \tilde{W}_n is the $n \times n$ matrix whose (i,j) th entry is $\hat{w}_{ijn} - E\hat{w}_{ijn}$.

Proof. Denote by M_n the $n \times n$ matrix whose entries are all $E\hat{w}_{11n}$.

Since

$$|E\hat{w}_{11n}| \leq E|w_{11}| I[|w_{11}| \geq \delta\sqrt{n}] \leq E|w_{11}|^{2/r} (\delta\sqrt{n})^{-\frac{2}{r}+1},$$

we obtain

$$\begin{aligned} & |\tilde{\lambda}_{\max}(n) - \hat{\lambda}_{\max}(n)| \\ & \leq n^{-2r} \left\{ 2 \sup_{\|a\|=1} |a' \hat{W}_n M_n a| + \sup_{\|a\|=1} |a' M_n^2 a| \right\} \\ & = n^{-2r} \left\{ 2 \sup_{\sum_{i=1}^p a_i^2 = 1} \left| \sum_{i=1}^p a_i \left| \sum_{j=1}^n a_j \hat{w}_{ijn} E\hat{w}_{11n} \right| \right| + n |E\hat{w}_{11n}|^2 \right. \\ & \quad \times \left. \sup_{\sum_{i=1}^p a_i^2 = 1} \left(\sum_{i=1}^p a_i \right)^2 \right\} \\ & \leq cn^{-2r} \left\{ n^{-\frac{1}{r}+1} \delta^{-\frac{2}{r}+1} \left(\sum_{i=1}^p \sum_{j=1}^n w_{ij}^2 \right)^{1/2} + n^{-\frac{2}{r}+3} \delta^{-\frac{4}{r}+2} \right\} \\ & \leq c \left\{ n^{-\frac{1}{r}(r-1)^2-1} \delta^{-\frac{2}{r}+1} \left(n^{-2r} \sum_{i=1}^p \sum_{j=1}^n w_{ij}^2 \right)^{1/2} \right. \\ & \quad \left. + n^{-\frac{2}{r}(r-1)^2-1} \delta^{-\frac{4}{r}+2} \right\} \rightarrow 0, \text{ a.s. if } r > 1. \end{aligned}$$

$$\leq c \left\{ n^{-\frac{1}{r}(r-1)^2 - r - \frac{2}{\delta} + 1} \left(n^{-2} \sum_{i=1}^p \sum_{j=1}^n w_{ij}^2 \right)^{1/2} + n^{-\frac{2}{r}(r-1)^2 - 1 - \frac{4}{\delta} + 2} \right\} \rightarrow 0, \text{ a.s. if } r \leq 1. \quad (2.2)$$

In proving the above result, we have used the facts that

$$n^{-2r} \sum_{i=1}^p \sum_{j=1}^n w_{ij}^2 \rightarrow 0, \text{ a.s. if } r > 1,$$

and

$$n^{-2} \sum_{i=1}^p \sum_{j=1}^n w_{ij}^2 \rightarrow E w_n^2 \text{ a.s. if } r \leq 1,$$

which can be seen from Marcinkiewicz strong law of large numbers.

Remark 1. Throughout this paper, we will use Lemmas 2.2 and 2.3 with $r = \frac{1}{2}$, and the requirement (b) is specified as

$$(b') \quad \delta \log n \rightarrow \infty, \text{ when } n \rightarrow \infty. \quad (2.3)$$

Remark 2. From Lemmas 2.2 and 2.3 we can easily see that

$$\lambda_{\max}(n) - \tilde{\lambda}_{\max}(n) \rightarrow 0, \text{ a.s. as } n \rightarrow \infty,$$

if the conditions of Theorem 1 hold, where $\lambda_{\max}(n)$ is the largest eigenvalue of $\frac{1}{n} W W' / n$

3. SOME RESULTS ON GRAPH THEORY

Given a sequence $(i_1, j_1, i_2, j_2, \dots, i_k, j_k)$, where i_1, \dots, i_k are integers in the set $\{1, \dots, p\}$, and j_1, \dots, j_k are integers in the set $\{1, \dots, n\}$, we define a directed multigraph as follows. We draw two parallel real lines, I-line and J-line. We plot i_1, \dots, i_k on the I-line and plot j_1, \dots, j_k on the J-line. These are vertices and they are split into two disjoint classes on the two parallel lines. So even if the two integers i_a and j_b are equal, they will not be the same vertex because i_a belongs to I-line and j_b belongs to J-line. But if $i_a = i_b$ (or $j_a = j_b$), we regard these two vertices identical. Edges of the directed bigraph will be the directed segments $\overrightarrow{i_1 j_1}$, $\overrightarrow{j_1 i_2}$, $\overrightarrow{i_2 j_2}$, $\overrightarrow{j_2 i_3}$, \dots , $\overrightarrow{i_k j_k}$, $\overrightarrow{j_k i_1}$. They are $2k$ in number and they should be regarded different from each other, even in the case when two edges have the same initials and ends.

Sometimes it is convenient to denote i_a by v_{2a-1} and j_a by v_{2a} . So the vertices of the graph are v_1, v_2, \dots, v_{2k} , and the edges are $v_1 v_2, v_2 v_3, \dots, v_{2k-1} v_{2k}, v_{2k} v_1$. Notice that when we write an edge as $v_a v_{a+1}$, we always mean that v_a is the initial vertex and v_{a+1} is the end vertex, the direction of the edge is from v_a to v_{a+1} .

When two edges $v_a v_{a+1}, v_b v_{b+1}$ have the same vertex sets, i.e. $\{v_a, v_{a+1}\} = \{v_b, v_{b+1}\}$, we cannot conclude that $v_a v_{a+1} = v_b v_{b+1}$, since $v_a v_{a+1} = v_b v_{b+1} \Leftrightarrow a = b$. When $\{v_a, v_{a+1}\} = \{v_b, v_{b+1}\}$, we say that the two edges coincide.

The graph we just constructed will be called a W-graph.

A W-graph will be called canonical, if $v_a \leq \max\{v_{a-2}, v_{a-4}, \dots\} + 1$, for each $a > 2$, and $v_1 = 1, v_2 = 1$.

In the following, we will get a bound for the number of canonical W-graphs.

In a canonical W-graph, an edge $v_{a-1}v_a$ ($a \geq 2$) will be called an innovation, if v_a does not occur in v_1, v_2, \dots, v_{a-1} . Suppose $v_{a-1}v_a$ is an innovation. If a is odd, $v_{a-1}v_a$ is called a row innovation, and if a is even, a column innovation. Note that v_1v_2 is a column innovation according to the above definition.

In a W-graph, an edge $v_{a-1}v_a$ ($a \geq 2$) is said to be single up to v_b ($b \geq a$), if there is no edge $v_{c-1}v_c$ with $1 < c \leq b$ such that $v_{c-1}v_c$ coincides with $v_{a-1}v_a$. An edge $v_{b-1}v_b$ ($b \geq 3$) will be called a T_3 -edge if there is an innovation $v_{a-1}v_a$, single up to v_{b-1} , and $v_{b-1}v_b$, $v_{a-1}v_a$ coincide.

An edge will be called a T_4 -edge if it is not an innovation and not a T_3 -edge.

A consecutive segment $v_a v_{a+1} \dots v_{b-1} v_b$ of the whole W-graph will be called a chain.

Lemma 3.1. Let $v_a v_{a+1} \dots v_c$ be a chain such that

- (1) $v_a v_{a+1}$ is single up to v_c ,
- (2) v_c has been visited by $v_1 v_2 \dots v_a$.

Then the chain contains at least one T_4 -edge.

Proof. When $c - a = 1$, evidently $v_a v_{a+1} = v_a v_c$ is an edge of T_4 .

We know that $v_{c-1}v_c$ must be a T_4 or a T_3 since it cannot be an innovation. If it is a T_4 , the proof is completed. If $v_{c-1}v_c$ is a T_3 , then there is a single innovation $v_{b-1}v_b$ coincident with $v_{c-1}v_c$ such that $b < c$.

Case 1. $b \geq a+1$. Since $v_c = v_{b-1}$ or $v_c = v_b$, then either $v_a v_{a+1} \dots v_{b-1}$ or $v_a v_{a+1} \dots v_b$ has the properties (1), (2) and is shorter than $v_a v_{a+1} \dots v_c$. By induction hypothesis $v_a v_{a+1} \dots v_b$ contains a T_4 -edge, but it is a part of the chain $v_a v_{a+1} \dots v_c$. So the original path contains a T_4 -edge.

Case 2. $b < a+1$, i.e. $v_{b-1}v_b$ is in the path $v_1v_2\dots v_a$, and then $v_{c-1} = (v_b \text{ or } v_{b-1})$ is visited by $v_1v_2\dots v_a$. Thus $v_av_{a+1}\dots v_{c-1}$ has the properties (1),(2). By induction, the lemma is proved.

Note: In a W-graph, the chain $v_1v_2\dots v_{a+1}$ determines completely whether the edge v_av_{a+1} is an innovation, a T_3 or a T_4 .

Lemma 3.2. If in the chain $v_1v_2\dots v_a$, there are s edges, each of which is single up to v_a and has a vertex equal to v_a and if t is the number of noncoincident T_4 edges in $v_1v_2\dots v_a$, then $s \leq t + 1$.

Proof. Let $v_{a_1-1}v_{a_1}, v_{a_2}v_{a_2+1}, \dots, v_{a_s}v_{a_s+1}$ be all the single edges such that $a_1 < a_2 < \dots < a_s < a$ and $v_{a_2}v_{a_2+1}, \dots, v_{a_s}v_{a_s+1}$ are single up to v_a , and

$$v_{a_1} = v_{a_2} = \dots = v_{a_s} = v_a.$$

Consider chains $L_2 = v_{a_2}v_{a_2+1} \dots v_{a_3}$, $L_3 = v_{a_3}v_{a_3+1} \dots v_{a_4}$, ..., $L_s = v_{a_s}v_{a_s+1} \dots v_a$. By Lemma 3.1, L_2 has a T_4 -edge E_2 . Let v_{b_3} be the first vertex in L_3 which belongs to L_2 . Then by Lemma 3.1

$v_{a_3}v_{a_3+1} \dots v_{b_3}$ contains an edge of T_4 and we denote it by E_3 .

Evidently, E_3 and E_2 are not coincident. Let v_{b_4} be the first vertex of L_4 which also belongs to $L_2 \cup L_3$. Then $v_{a_4}v_{a_4+1} \dots v_{b_4}$ has an edge of T_4 , by Lemma 3.1. Let it be denoted by E_4 . Evidently no two of E_2, E_3, E_4 are coincident. Continue this procedure. Finally, we get $s-1$ edges of T_4 , which are not coincident with each other. So $s-1 \leq t$, $s \leq t+1$.

A T_3 edge v_av_{a+1} is called regular, if there is more than one innovation with a vertex equal to v_a and single up to v_a .

Lemma 3.3. In any W-graph there is a mapping ϕ from regular T_3 edges to T_4 edges such that for any T_4 edge E , there are at most two regular T_3 edges whose ϕ image is E .

Proof. Define ϕ as follows.

Let $v_{a_1} v_{a_1+1}, v_{a_2} v_{a_2+1}, \dots, v_{a_s} v_{a_s+1}$ be the set of all innovations single up to v_a such that

- (1) $v_{a_2} = v_{a_3} = \dots = v_{a_s} = v_a$
- (2) $v_{a_1} = v_a$ or $v_{a_1+1} = v_a$
- (3) $a_1 < a_2 < \dots < a_s < a_{s+1} = a$.

We note that there is at most one innovation inward to v_a and if there is such an innovation, it must be the foremost one among innovations with a vertex v_a .

In this W-graph, $v_a v_{a+1}$ must coincide with one of $v_{a_1} v_{a_1+1}, \dots, v_{a_s} v_{a_s+1}$. Let that one be $v_{a^*} v_{a^*+1}$. Also, let

$$v = v(a) = \begin{cases} i + 1, & \text{if } a^* = a_i, \text{ for } i = 1, \dots, s-1, \\ s, & \text{if } a^* = a_s. \end{cases}$$

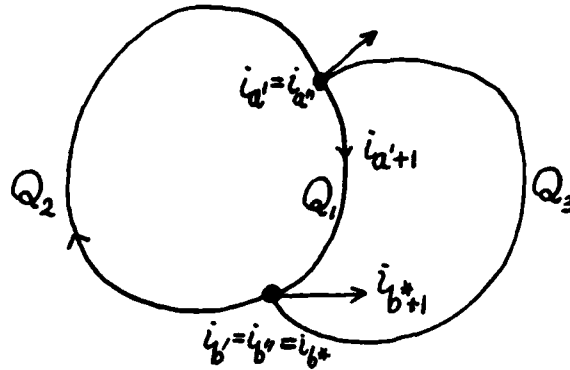
By Lemma 3.1, in the path $C_a = v_{a_1} v_{a_1+1} \dots v_{a_v+1}$, there is at least one edge of T_4 . Let the first one be E_a . Define $\phi(v_a v_{a+1}) = E_a$.

At first we prove that if $v_a v_{a+1}, v_b v_{b+1}$ are two regular T_3 edges and $v_a \neq v_b$ then $E_a \neq E_b$ (E_a, E_b may coincide). Suppose $C_a = v_a v_{a'+1} \dots v_{a''}$, $C_b = v_b v_{b'+1} \dots v_{b''}$. Then, we have the following possibilities to consider:

- (1) $a' < a'' < b' < b''$,
- (2) $a' < b' < a'' < b''$,
- (3) $a' < b' < b'' < a''$.

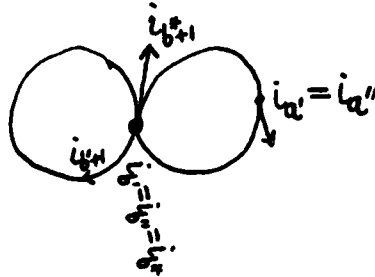
For case (1), $E_a \neq E_b$ is evident.

Consider case (2). C_a, C_b are divided into three parts as $Q_1 = v_a v_{a'+1} \dots v_{b'}$, $Q_2 = v_{b'} v_{b'+1} \dots v_{a''}$, $Q_3 = v_{a''} v_{a''+1} \dots v_{b''}$ as given in the following diagram.



It is enough to show that E_a is in Q_1 , E_b is in $Q_2 \cup Q_3$. By definition, there is an innovation $v_{b*}v_{b*+1}$ with $b* < b'$, $v_{b*} = v_{b'} = v_b$, single up to v_b . If $b* \leq a'$, by Lemma 3.1, Q_1 contains a T_4 -edge. If $a' < b* < b'$, we can consider the chain $v_{b*}v_{b*+1} \dots v_b$; it is part of Q_1 and it contains a T_4 -edge by Lemma 3.1. So $E_a \in Q_1$. It is obvious that E_b is in $Q_2 \cup Q_3$. Thus $E_a \neq E_b$.

Consider (3). As before, we can show that E_a is in $v_a, v_{a'+1} \dots v_b$, and E_b is in $v_b, v_{b'+1} \dots v_b''$.



Then we consider the case $v_a = v_b$ and $a < b$. Now $v_a v_{a+1}$, $v_b v_{b+1}$ both are regular T_3 edges and so they coincide with single innovations $v_{a*}v_{a*+1}$ and $v_{b*}v_{b*+1}$. $v_{a*}v_{a*+1}$ and $v_{b*}v_{b*+1}$ cannot coincide. So $E_a \neq E_b$ except $v_{a*}v_{a*+1}$, $v_{b*}v_{b*+1}$ are the last two single innovations. In the last case $\phi^{-1}(E_a)$ has cardinal 2.

At last we get that the mapping $\phi: v_a v_{a+1} \longrightarrow E_a$ for regular T_3 's has the property that $\phi^{-1}(E_a)$ has at most two edges of T_3 . The proof of Lemma 3.3 is completed.

4. LIMIT OF THE LARGEST EIGENVALUE

We now prove the main result of our paper.

Theorem 3.1. Let $\{w_{ij} : i, j = 1, 2, \dots\}$ be an infinite matrix of iid random variables, $Ew_{11} = 0$ and $Ew_{11}^4 < \infty$. If $\lambda_{\max}(n)$ denotes the largest eigenvalue of the matrix $\frac{1}{n} W_n W_n'$, here W_n denotes the $p \times n$ random matrix $\{w_{ij}; i = 1, \dots, p; j = 1, \dots, n\}$, then

$$\lim_{n \rightarrow \infty} \lambda_{\max}(n) = (1 + \sqrt{y})^2 Ew_{11}^2 \text{ a.s.}$$

as $n \rightarrow \infty$, $p \rightarrow \infty$ and $p/n \rightarrow y$.

Proof. Without loss of generality, we assume that $Ew_{11}^2 = 1$. We only have to prove that $\overline{\lim} \lambda_{\max}(n) \leq (1 + \sqrt{y})^2$ a.s. . But, according to Remark 2 of Section 2, it is sufficient to show that

$$\overline{\lim} \tilde{\lambda}_{\max}(n) \leq (1 + \sqrt{y})^2 \text{ a.s. .}$$

In other words, we assume in the sequel that

- (1) $|w_{ij}| < \delta\sqrt{n}$,
- (2) $Ew_{ij} = 0$,
- (3) $Ew_{ij}^2 \leq 1$,
- (4) $E|w_{ij}|^l \leq (\delta\sqrt{n})^{l-1}$, for $l \geq 2$,
- (5) $E|w_{ij}|^l \leq c(\delta\sqrt{n})^{l-3}$, for $l \geq 3$.

Now, choose $z > (1 + \sqrt{y})^2$ arbitrarily. We will now show that

$$(6) \quad \sum_{n=1}^{\infty} E \left(\frac{\lambda_{\max}(n)}{z} \right)^k < \infty$$

where $k = k_n$ satisfies

- (7) $k_n / \log n \rightarrow \infty$,
- (8) $\delta^{1/6} k_n / \log n \rightarrow 0$.

We have

$$\begin{aligned} E[\lambda_{\max}(n)]^k &\leq E \operatorname{tr} \left(\frac{1}{n} W W^T \right)^k = n^{-k} E \operatorname{tr} (W W^T)^k \\ &= n^{-k} \sum E w_{i_1 j_1} w_{i_2 j_1} w_{i_2 j_2} \cdots w_{i_k j_k} w_{i_1 j_k}. \end{aligned}$$

Here the summation is taken in such a way that i_1, \dots, i_k run over all integers in $\{1, \dots, p\}$, and j_1, \dots, j_k run over all integers in $\{1, \dots, n\}$.

The above sum can be split in the following form.

$$E \operatorname{tr} \left(\frac{1}{n} W W^T \right)^k = n^{-k} \sum' \sum'' \sum''' E w_{i_1 j_1} w_{i_2 j_1} w_{i_2 j_2} \cdots w_{i_k j_k} w_{i_1 j_k},$$

here

\sum' - summation for different arrangements of four different types of elements at the $2k$ different positions.

\sum'' - summation for different canonical graphs Γ with given arrangement of the four types for $2k$ positions.

\sum''' - summation of $E w_{i_1 j_1} w_{i_2 j_1} \cdots w_{i_1 j_k}$ for which the graph is isomorphic to the given canonical graph.

Let r denote the number of row innovations, ℓ denote the number of T_3 edges. Then there are $\ell - r$ column innovations and $(2k - 2\ell)T_4$ edges and so \sum' is bounded by $\sum_{\ell=1}^k \sum_{r=1}^{\ell} \binom{k}{r} \binom{k}{\ell-r} \binom{2k-\ell}{\ell}$. Since every row innovation leads to a free i -index and every column innovation leads to a free j -index except the first column innovation $v_1 v_2$ which leads to an i -index and a j one. We know that \sum''' is bounded by $p^{r+1} n^{\ell-r}$.

To bound \sum'' , let t denote the number of noncoincident T_4 edges. By definition, each innovation in a canonical W -graph is uniquely determined by the chain before it and each nonregular T_3 edge is so

done as innovations. If $t = 0$, i.e. $\ell = k$, then by Lemma 3.3, there are no T_4 edge and regular T_3 edge, so that \sum'' is only one summand. For $t \geq 1$, since each T_4 edge in a canonical W -graph must be one of the k^2 elements in the $k \times k$ matrix (w_{ij}) $i \leq k, j \leq k$, all the possibilities that $2k - 2\ell$ T_4 edges may take are less than $\binom{k^2}{t} t^{2k-2\ell}$. By Lemmas 3.2 and 3.3, all the possibilities that all the regular T_3 edges take is not more than $(t+1)^{4k-4\ell}$. Hence \sum'' is bounded by $\binom{k^2}{t} t^{2k-2\ell} (t+1)^{4k-4\ell} \leq k^{2t} (t+1)^{6k-6\ell}$.

Finally, we bound the expectation $E w_{i_1 j_1} w_{i_2 j_1} \cdots w_{i_k j_k} w_{i_1 j_k}$. If $t = 0$, each expectation is $(E w_{11}^2)^\ell \leq 1$. For $t \geq 1$, let μ denote the number of innovations which coincide with at least one T_4 edge and let n_i denote the number of T_4 edges which coincide with the i -th such innovation, $i = 1, 2, \dots, \mu$, respectively.

Let m_j be the number of T_4 edges which coincide with each other but not with any innovation, $j = 1, 2, \dots, t-\mu$. Then we have

$$E w_{i_1 j_1} w_{i_2 j_1} \cdots w_{i_k j_k} w_{i_1 j_k} = (E w_{11}^2)^{\ell-\mu} \prod_{i=1}^{\mu} (E w_{11}^{n_i+2}) \prod_{j=1}^{t-\mu} E w_{11}^{m_j}$$

where $2(\ell-\mu) + \sum_{i=1}^{\mu} (n_i + 2) + \sum_{j=1}^{t-\mu} m_j = 2k$ and $\mu \leq t$. By (3), (4) and

(5) we have

$$\begin{aligned} |E w_{i_1 j_1} w_{i_2 j_1} \cdots w_{i_k j_k} w_{i_1 j_k}| &\leq c^{\mu(\delta\sqrt{n})} \prod_{i=1}^{\mu} (n_i - 1) + \sum_{j=1}^{t-\mu} (m_j - 1) \\ &= c^{\mu(\delta\sqrt{n})} 2k - 2\ell - t \leq k^t (\delta\sqrt{n})^{2k-2\ell-t} \end{aligned}$$

when k is large enough.

By the above argument, we obtain

$$E \operatorname{tr} \left(\frac{1}{n} W_n W_n^T \right)^k \leq n^{-k} \sum_{\ell=1}^k \sum_{r=1}^{\ell} \binom{k}{r} \binom{k}{\ell-r} \binom{2k-\ell}{\ell} p^{r+1} n^{\ell-r} \left[\sum_{t=0}^{2k-2\ell} k^{2t} (t+1)^{6k-6\ell} \right. \\ \left. k^t (\delta \sqrt{n})^{2k-2\ell-t} \right]$$

$$\leq p \sum_{\ell=1}^k \sum_{r=1}^{\ell} \binom{k}{r} \binom{k}{\ell-r} \binom{2k-\ell}{\ell} \left(\frac{p}{n} \right)^r \left[\sum_{t=0}^{2k-2} k^{3t} (t+1)^{6k-6\ell} (\delta \sqrt{n})^{-t} \right] \delta^{2(k-\ell)}$$

Using the elementary inequality $a^{-t} (t+1)^b \leq \left(\frac{b}{\log a} \right)^b$ for $a > 1$, $b > 0$,

$t > 0$, we have

$$\sum_{t=0}^{2k-2\ell} k^{3t} (t+1)^{6k-6\ell} (\delta \sqrt{n})^{-t} \leq 2k \left(\frac{6k-6\ell}{\log \frac{\delta \sqrt{n}}{k^3}} \right)^{6k-6\ell} \\ \leq 2k \left(\frac{6k}{\frac{1}{2} \log n + \log \delta - 3 \log k} \right)^{6(k-\ell)} \\ \leq 2k \left(\frac{18k}{\log n} \right)^{6(k-\ell)} \quad \text{when } n \text{ is large enough.}$$

Using Lemma 2.1, we get for all large n

$$E \operatorname{tr} \left(\frac{1}{n} W_n W_n^T \right)^k \\ \leq 2kp \sum_{\ell=1}^k \sum_{r=1}^{\ell} (1+\sqrt{\delta})^{2k} \binom{k}{\ell} \binom{2\ell}{2r} \left(\frac{p}{n} \right)^r \left[\frac{18\delta^{1/6} k}{\log n} \right]^{6(k-\ell)} \\ \leq 2kp (1+\sqrt{\delta})^{2k} \left[(1+\sqrt{p/n})^2 + \left[\frac{18\delta^{1/6} k}{\log n} \right]^6 \right]^k \\ = (2kp)^{1/k} (1+\sqrt{\delta})^2 \left((1+\sqrt{p/n})^2 + \left[\frac{18\delta^{1/6} k}{\log n} \right]^6 \right)^k \\ \leq \eta^k,$$

where η is a constant satisfying $(1+\sqrt{y})^2 < \eta < 2$. Here the last inequality follows from the following facts:

- (a) $(2kp)^{1/k} \rightarrow 1$, because $k/\log n \rightarrow \infty$ and $p/n \rightarrow y \in (0, \infty)$
- (b) $(1+\sqrt{\delta})^2 \rightarrow 1$, because $\delta \rightarrow 0$
- (c) $(1+\sqrt{p/n})^2 \rightarrow (1+\sqrt{y})^2$, because $p/n \rightarrow y \in (0, \infty)$
- (d) $\left(\frac{18\delta^{1/6} k}{\log n} \right)^6 \rightarrow 0$, because $\delta^{1/6} k / \log n \rightarrow 0$.

This leads to (6) since $k/\log n \rightarrow \infty$ and the proof is thus complete.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 84-44		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 85-0006	
6a. NAME OF PERFORMING ORGANIZATION University of Pittsburgh	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6c. ADDRESS (City, State and ZIP Code) Center for Multivariate Analysis 515 Thackeray Hall, Pittsburgh PA 15260		7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-82-K-0001	
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448		10. SOURCE OF FUNDING NOS	
		PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304
		TASK NO. A5	WORK UNIT NO.
11. TITLE (Include Security Classification) ON LIMIT OF THE LARGEST EIGENVALUE OF THE LARGE DIMENSIONAL SAMPLE COVARIANCE MATRIX			
12. PERSONAL AUTHOR(S) Y.Q. Yin*, Z.D. Bai* and P.R. Krishnaiah			
13a. TYPE OF REPORT Technical	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) OCT 84	15. PAGE COUNT 16
16. SUPPLEMENTARY NOTATION *On leave of absence from China University of Science and Technology.			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB GR	
		Largest eigenvalue; sample covariance matrix; large dimensional random matrices; limit.	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) In this paper, the authors showed that the largest eigenvalue of the sample covariance matrix tends to a limit under certain conditions when both the number of variables and the sample limit size tend to infinity. The above result is proved under the mild restriction that the fourth moment of the elements of the sample sums of squares and cross products (SP) matrix exist.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff		22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027	22c. OFFICE SYMBOL NM

END

FILMED

3-85

DTIC